General Game Playing with Stochastic CSP - one page abstract

General Game Playing with Stochastic CSP

Frédéric Koriche, Sylvain Lagrue, Éric Piette, and Sébastien Tabary

Université Lille-Nord de France CRIL - CNRS UMR 8188 Artois, F-62307 Lens
{koriche,lagrue,epiette,tabary}@cril.fr
http://www.cril.univ-artois.fr/~{koriche,lagrue,epiette,tabary}

Abstract. [1] The aim of *General Game Playing* (GGP) is to devise game playing algorithms which are not dedicated to a particular strategic game, but are general enough to effectively play a wide variety of games. A tournament is held every year by AAAI, in which artificial game players are supplied the rules of arbitrary new games and, without human intervention, have to play these game optimally. Games rules are described in a declarative representation language, called GDL for *Game Description Language*. The lastest version of this language is expressive enough to describe finite multi-player games with uncertain and incomplete information. GGP algorithms include, among others, answer set programming methods, automated construction of evaluation functions, and Monte-Carlo methods such as Upper Confidence bounds for Trees (UCT). Beyond its play value, GGP offers a rigorous setting for modeling and analyzing sequential decision-making algorithms in multi-agent environments.

By providing a declarative approach for representing and solving combinatorial problems, Constraint Programming appears as a promising technology to address the GGP challenge. Currently, several constraint-based formalisms have already been proposed to model and solve games; they include notably *Quantified* CSP, *Strategic* CSP and *Constraint Games*. Most of these formalisms are, however, restricted to deterministic, perfect information games: during each round of the game, players have full access to the current state and their actions have deterministic effects. This paper focuses on *stochastic* games, with chance events, using the framework of stochastic constraint networks.

More precisely, we study a fragment of the Stochastic Constraint Satisfaction Problem (SCSP), that captures GDL games with uncertain (but complete) information. Interestingly, the SCSP for this class of games can be decomposed into a sequence of μSCSPs (a.k.a one-stage stochastic constraint satisfaction problems). Based on this decomposition, we propose a sequential decision-making algorithm, MAC-UCB, that combines the MAC algorithm (*Maintaining Arc Consistency*) for solving each μSCSP, and the multi-armed bandits *Upper Confidence Bound* (UCB) method for approximating the expected utility of strategies. We show that in practice MAC-UCB significantly outperforms (the multi-player version of) UCT, which is the state-of-the-art GGP algorithm for stochastic games. MAC-UCB also dominates FC-UCB, a variant where MAC is replaced with the Forward Checking (FC) algorithm. Such conclusions are drawn from comparing the performance of these algorithms, using extensive experiments (about $1,350,000$ face-offs) over a wide range of GDL games.

[1] This paper has been published in Constraints 21(1), the Journal Fast Track issue of CP'15.